CASE TYPE PREDICTION AND MITIGATION OF READER'S COMPLEXITY IN LEGAL DOMAIN USING DEEP LEARNING TECHNIQUES

A SYNOPSIS

Submitted by

V.VAISSNAVE
(Reg. No. 201907176)

in partial fulfillment for award of the degree

of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SCHOOL OF COMPUTING
KALASALINGAM ACADEMY OF RESEARCH AND EDUCATION
(DEEMED TO BE UNIVERSITY)
ANAND NAGAR, KRISHNANKOIL – 626126
TAMILNADU, INDIA

NOVEMBER 2022

1. INTRODUCTION

The Indian judicial system is substantially based on the English common law system in which judges produce law through decisions, orders, and judgments. It has established a federal structure with central government and state governments working together. The Indian judicial system is one of the world's largest judiciary systems and it deals with a significant number of legal cases that are growing at a rapid rate. Legal texts greatly influence the interpretation and application of the law. Indian law documents are being digitized. Hence, the status of cases are available to petitioners and lawyers can consult legal data online. Due to the increasing volume of legal data, it is being computerized into complex forms, which in turn makes it difficult for lawyers to pertain. Research on document summarization, text classification, case outcome prediction, keyword extraction and audio transcription are happening to ease the overhead for legal entities to process legal documents.

The process of detecting the positive or negative sentiments in a data is referred as Sentiment Analysis (SA) or opinion mining [1]. It is used in natural language processing to extract the sentiments from social data and also used in business to gain a better understanding of their customers. Opinion mining uses the polarity of a text data to classify it as positive, negative or neutral. It even detects the emotions and intentions in several domains [2]. Legal domain is in need of such an approach, which can be utilized for case outcome prediction to favor the public and lawyers.

Legal decision necessitates reviewing numerous legal documents. Extracting relevant information from legal text documents manually is a tedious and time-consuming task. Various legal elements influence the result of a legal case, and these factors might vary substantially depending on the case type [3]. Such extraction will be beneficial for legal readers and also it will help in the prediction of success rate of judges.

Document Classification is the act of labeling the documents into multiple categories.

Since law documents are of multiple pages, it will be useful for the legal readers if documents are categorized according to the content type [4]. Document can be labeled manually which is a hectic task and will not be accurate. Automated labeling of legal documents using deep learning techniques are evolving [5]. Existing document classification needs more improvement for precise classification of legal documents.

Text summarization is one of the most useful techniques to reduce the content in legal document. Data summarization can aid in data interpretation and decision-making [6]. A good summary should preserve the essential substance from the enormous amount of documents. The legal text summarizing extracts the most useful data of a document and condenses them for a lawyer for their reference. Such short formats of documents must be automatically created by machines without human intervention in an accurate manner [7].

The proposed research addresses these complexities and provides an efficient case outcome prediction model and also readers complexity mitigation techniques like summarization, information extraction and categorizing the judgment text into many distinct categories are implemented. The proposed research is an integration of several novel techniques and goes through multiple phases to provide accurate data for legal readers.

2. LITERATURE SURVEY

Computerized judgment text has gained attention throughout the last decade. Several methodologies have been introduced to resolve issues in judgment text analysis and to minimize the complexity for the legal text readers.

2.1 Survey on Sentiment analysis

Sentiment analysis is a well-known opinion mining technique that has been employed in a variety of developing fields. Jianqiang et al. [8] used the co-occurrence statistical

features and context semantic properties of tokens in the tweeter dataset to estimate the n-gram values. Based on these input values, the CNN model is used to classify the sentiment text. Hyun et al. [9] employed a Target-dependent convolutional neural network (TCNN) for Target-level sentiment analysis (TLSA). TLSA is a classification technique to extract sentiments from targets in text. A modified LSTM was created by Ma et al. [10]. A separate output gate for storing word measure memory and concept measure input has been added to their model. It established RNN's mathematical work. Hu et al. [11] developed a short-text word collection and used it in conjunction with the CNN model. Though it has been used widely, it is new to the field of legal domain. This study utilized the benefits of sentiment analysis in the legal arena to determine whether the judgment text was an appeal or a dismissal case.

2.2 Survey on Information Extraction

The court legal documents include multiple information like petitioner name, wealth of information, hearing dates, case numbers, attorneys for both sides, etc. Manual extraction of such information is time consuming and prone to human errors, since all the cover sheets must be read to locate the necessary information. Many researches are being conducted in automated extraction of such information using Natural Language Processing (NLP) and deep learning techniques.

Karim et al. [12] has developed a Named Entity Recognition (NER) System for the Bangla language. NER uses Bidirectional Long Short Term Memory (BLSTM) and word embedding techniques. The extracted information falls under four categories: person names, places, organization names and object entities. In the Chinese legal area, Li et al. [13] established a cognitive computing framework to address the issues of judicial reasoning and knowledge learning. No previous studies have been able to predict a forthcoming judgement using legal filings. The suggested methodology employed by Liu & Chen [14] uses text mining algorithms to extract features from precedents but semantic comprehension of text inexact.

2.3 Survey on Document Summarization

Document summarization is the process of condensing a larger document into a shorter form. Since judgment text contains lots of information, a summarization process is adopted to reduce the document content. Few of such techniques are discussed below. Anand and Wagh [16] used a generic Neural Network (NN) based method for summarizing the Indian legal judgment documents. FirdausSolihin et al [15] present the word and phrase embedding are used to capture the semantics of the text using two NN frameworks. This approach does not rely on handcrafted characteristics or domain specific expertise and it is not limited to certain domains. The problem of not having access to labeled data is resolved by assigning scores or classes to each sentence in the training set. Tiwari and Dembla [17] propose a lexical chain computation method, which uses MAX taggers for statistical calculations and parts of speech tagging.

Zhang et al. [18] present a unique generative method based on the convolution seq2seq framework. In comparison to the classic RNN seq2seq paradigm, a hierarchical CNN design is extremely effective. A copying technique is used for managing unseen/rare terms in these models. They also incorporate a concurrent hierarchical attention approach for modeling essential sentences and keywords at the same time. Rahman and Siddiqui [19] developed an abstractive text summarization method that manually constructs a summary from lengthy texts but these techniques were not efficient for larger documents.

2.4 Survey on Document classification

The practice of giving categories or classes to documents in order to make them easier to search, manage or analyses is known as document categorization. Embracing document classification in the legal domain is extremely useful for the judges and other document readers. Several approaches are used for document classification, a Long Short Term Memory (LSTM) based approach to classify Chinese dataset was created by Wan et al., [20] uses long-term dependencies to determine the relationship between

entities. Iftikhar et al. [21] utilized the Lahore High Court (LHC) dataset to create an intelligent system that can locate and automatically classify important information from criminal legal judgments. The classified contents were labeled into nine elements such as FIR No., location, person name, date, and so on. Yan G, et al. [22] used the Chinese Supreme Court database to calculate the document similarity of Chinese decision documents. Initially, the existing language and taxonomy of public judgment documents were used to create a top-level ontology. The document classification was effective but it was a time consuming process.

From the existing literature survey, it was found that there is a need for highly precise information extraction and case outcome prediction techniques, better scheme for document summarization and classification of judgmental text. Hence, this research focuses on providing methodologies to overcome these existing challenges in the legal domain.

3. RESEARCH PROBLEM

According to the conducted research, the following are the common and persisting issues in the legal arena.

- Legal document case outcome prediction is a tedious task. So, an efficient technique is necessary for faultless sentiment classification of a legal document as appeal or dismissal.

- Extraction of key features in the legal arena is difficult due to its sensitivity. The loss of an important keyword is dropping the performance of existing techniques.

- Current text summarization techniques are not efficient on legal document summarization; they do not provide optimal solutions and lack in evaluation metrics for legal document summaries.

- Noise free text classification technique is necessary to avoid the misclassification

of legal data.

4. RESEARCH OBJECTIVE

Our major research focus is to develop an effective case outcome prediction method and to use mitigation measures to reduce the reading complexity of legal documents. This leads us to have the following as our research objectives.

- **Objective 1:** To develop a faultless deep learning based opinion mining technique for precise case outcome prediction.
- **Objective 2:** To develop a novel keyword extraction technique using named entity recognition and part of speech tagging and also to implement a success rate predictor for finding the successful cases of judges and the associated case type.
- **Objective 3:** To develop an automated optimal deep learning based legal text summarization (ODL-LTS) approach and to opt for evolving summary evaluation metrics to test the performance of the developed approach.
- **Objective 4:** To develop a flawless multi-class text classification approach using Continuous Bag of Word model (CBOW) and LSTM to categorize data into multiple independent classes.

5. PROPOSED METHODOLOGY

The proposed research begins with a deep learning based opinion mining (DOM) approach which determines the main structure of the case as appeal (positive) or dismiss (negative). Word embedding techniques and deep learning models are used to precisely find the sentiment. The appeal case acquired from DOM is further utilized to determine the judge's success rate. An automated Success Rate Predictor (SRP) is proposed to determine the success rate of judges as well as to predict the case type of the legal document.The SRP processes the data received from Information Extractor (IE). IE consists of an unambiguous keyword extraction technique named as Named Entity

Recognition (NER) and Parts of speech tagging for information extraction. A novel Key Extracting technique (KeyExt) is proposed to extract the information from the NER list. The law information, judge name and case outcome are extracted by IE, such information is used by the SRP to find the case type. The case information stored in the database is used to compute the success rate of judges, which is helpful for lawyers to file their case. Further to minimize the intricacy of document reference for the legal readers, a document summarization approach named Optimal DL based Legal Text Summarization (ODL-LTS) is proposed.

ODL-LTS is an automated approach, which generates the labeled training data to segment the legal text by the use of domain knowledge. ODL-LTS decreases the overhead of reviewing the large legal documents into short significant information. Summarized text will give an overview of the case file but there will be a need for a lawyer to refer to a part of the judgment document in a certain category. It will be difficult to search the needed category from a large legal document. To mitigate this issue, a technique to categorize the legal document into multiple classes is proposed and named as Legal Data Classifier (LDC). LDC classifies data into eight categories. The output produced by LDC is obligatory and inevitable information, which are intended to generate information useful for lawyers. The proposed research addresses the discussed issues in the legal domain using the methodologies depicted in Figure 1. The Indian Supreme Court judgment text is used as a real-time dataset for sentiment analysis. The datasets were gathered from legal websites and divided into 5000 positive and 5000 negative texts. Based on the judgment text, the sentiment will be either “positive” or "negative".

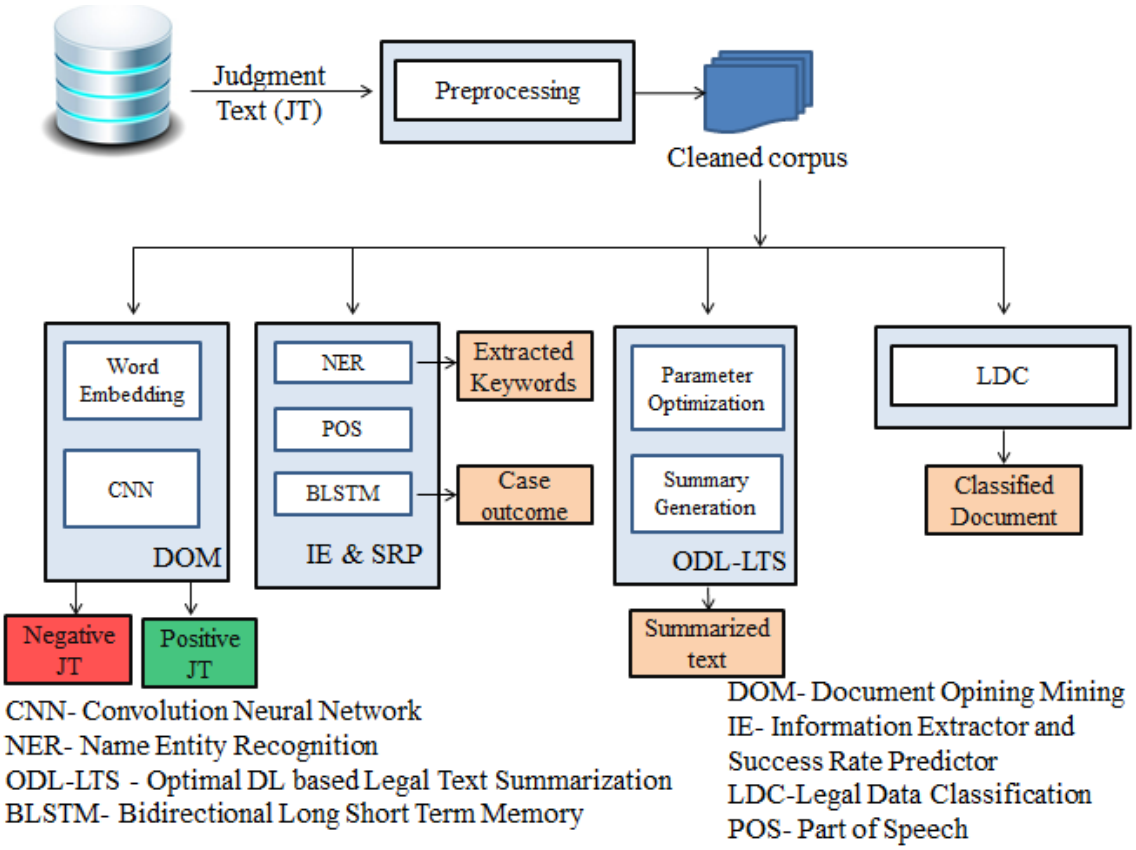

Figure 1: Proposed Research Framework

5.1 Deep learning based Opinion Mining

Deep learning based Opinion Mining (DOM) is proposed for case outcome prediction. The judgment text is generally classified as positive or negative i.e. case appeal or case dismissal. The judgement text is pre-processed and then the cleaned corpus is fed into Global Vector for word representation (GloVe). Convolution Neural network follows GloVe to classify the correct sentiment. The architecture of DOM is shown in Figure 2.

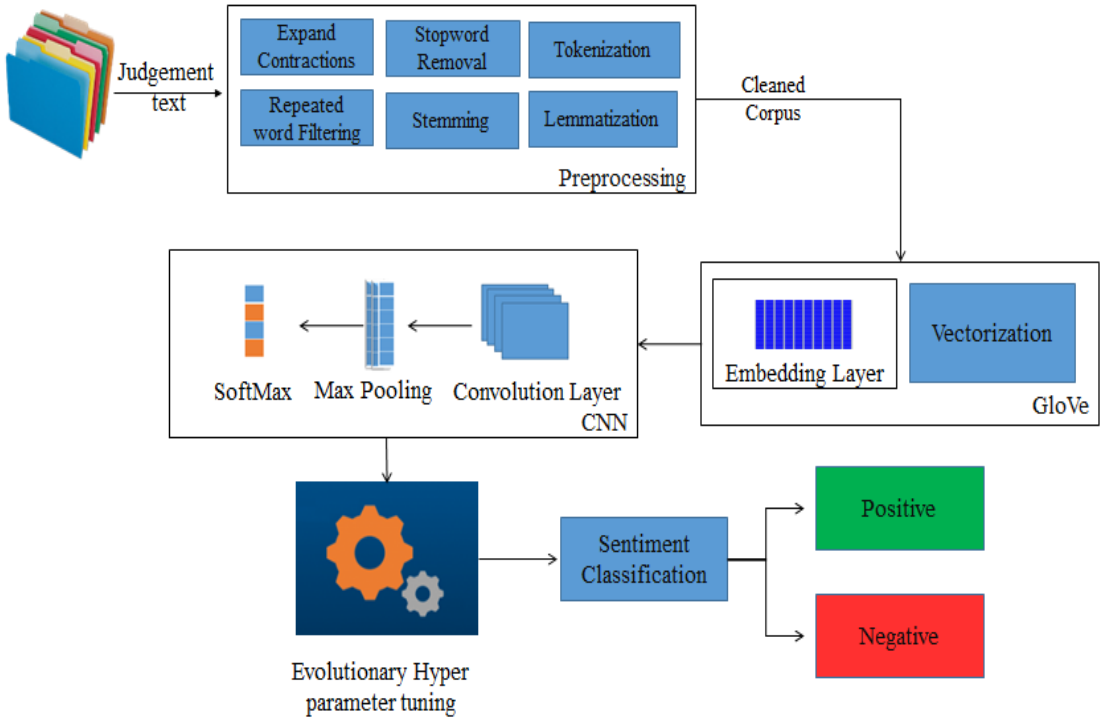

Figure 2: Proposed Document Opinion Mining

5.1.1 Data Pre-processing

Data processing is applied to produce a cleaned corpus. Expansion of contractions is the initial step in text standardization and it is achieved by eliminating the contractions. Contractions is a shortened word produced by combinations of words dropping letters and substituting an apostrophe. Stop words are frequently used words like a, an , the, etc. and removing such frequent but low level information will be helpful in focusing the key information. Tokenization is the process of breaking the raw data into words or sentences, which will be useful in better understanding of context. Rephrase text is changing the text so that it will be easy to identify similar text. Stemming reduces the words to root stem by removing the prefix or suffix of a word. It can sometimes stem from unwanted words, leading to lemmatization, a process of stemming to a meaningful word.

5.1.2 Global vector and Convolution Neural network

The cleaned corpus is fed into Global Vector for word representation (GloVe). GloVe produces a word vector by combining the local statistics with global statistics for matrix factorization, where local statistics gives context information of words and global statistics are word co-occurrence. The lexicons in legal documents are represented as a real valued vector in a high-dimensional space. Words with similar meanings in the vector space have similar representations. From the embedding matrix, the embedding layer maps the words to their embedding vectors. The matrix produced by GloVe is given to the Convolution Neural Network (CNN). Convolution layer detects the features; it uses a filter to glide over the matrix which is of fixed size. The max pooling identifies the key features which are used for precise classification of sentiments. Finally, softmax normalizes the values produced by max pooling. Hyper parameter tuning is used to identify the key feature for precise sentiment recognition using back propagation. The tuning is done based on the learning rate, number of epochs, activation function and dropouts. After the training phase, the resultant model after back propagation is used for testing purposes to determine the opinion on judgement text.

5.1.3 Experimental analysis

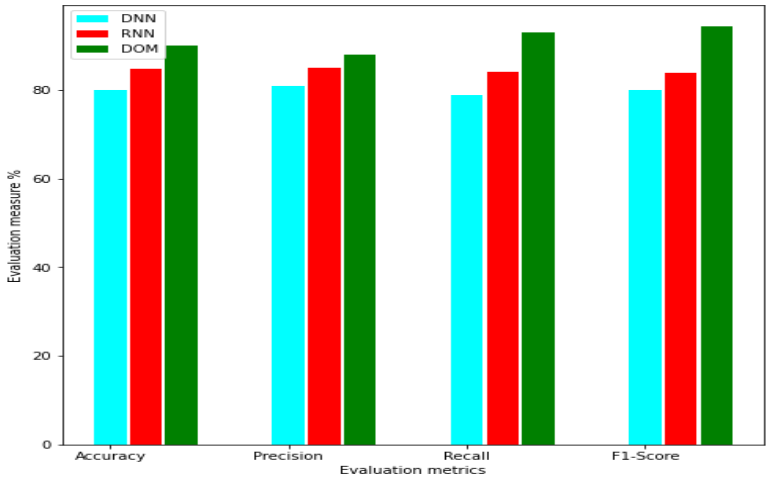

Figure 3: Performance measure of DOM

The proposed DOM was implemented using python. DOM was compared with models like Deep Neural Network (DNN) and Recurrent Neural Network (RNN). The performance measures of DOM based on metrics such as, accuracy, precision, recall are shown in Figure 3. The amount of words in the vocabulary has a major role in the model's effectiveness. Starting with 5000 words in the vocabulary list, the words are reduced by 1000 as shown in Table 1. DOM performance was optimal when the vocabulary contains 2,000 words. Dropout was used in DOM to improve the model's classification accuracy. The results are best when the dropout value is 0.4.

Table 1: DOM for different vocabulary size

Number of words	Accuracy	Precision	Recall	F1 score
5000	93.15	94	93	93
4000	93	91.6	92.6	92.5
3000	93.6	92.2	91.8	93.2
2000	**93.1**	**93**	**94.8**	**93.6**
1000	93.4	93.4	93.9	93.1
500	93.1	93	92.9	93.2

5.2 Information Extraction and Success rate prediction

The judgment text contains lots of information making it a tedious process to extract the information manually and such extraction will be open to errors. An efficient and automated information extraction technique named IE is proposed for automated data extraction. In IE, the cleaned corpus judgment file is fed into the transformation tagger. The tagged corpus file is processed by the Named Entity Recognition (NER) which extracts all the information as a list. The list is passed to the proposed Keyword Extractor (KeyExt), which extracts the necessary features required for case outcome prediction. The result of KeyExt is further fed into a trained Success Rate Predictor (SRP) for predicting the case type of a judgement text and then judge’s success rate is predicted for different case types as shown in Figure 4.

5.2.1 Information Extraction

Transformation-based tagging is a rule-based method for automatically tagging Parts Of Speech (POS) to the provided text. It changes from one state to another by using transformation rules, allowing us to have linguistic knowledge in a legible way. The transformation based tagging produces the tagged corpus. NER processes the tagged corpus and locates the information like name, section, location, case outcome, etc. for extraction. The NER obtains individual names from texts of judgments. KeyExt gets all the keywords and extracts the judge's name using the data scraped from a legal website. The judge name and case outcome information is taken out to find the success and failure rate of the cases for a particular judge, which will be used as a reference for lawyers.

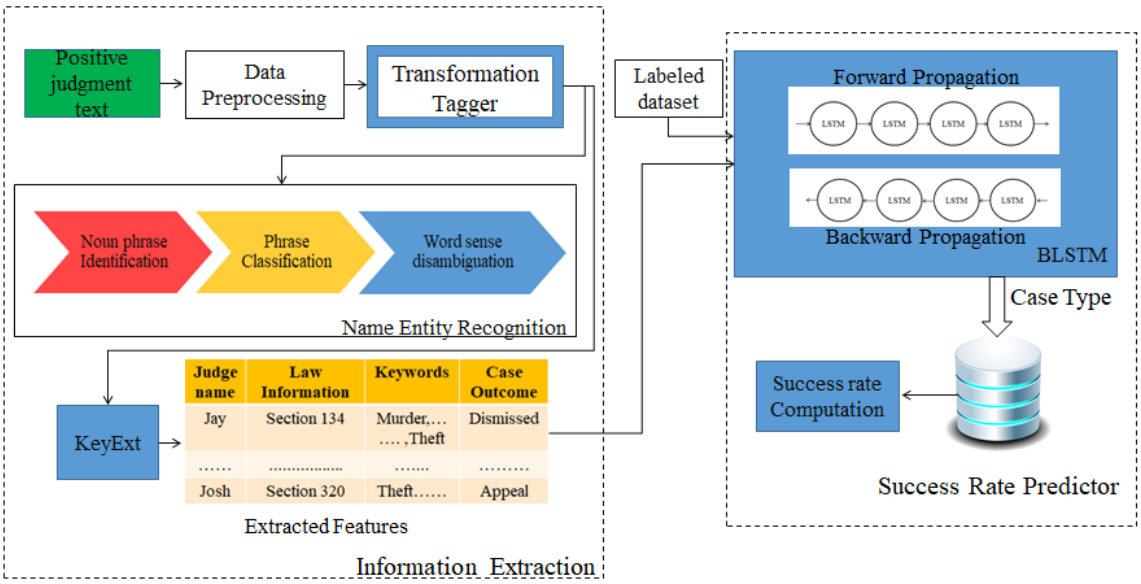

Figure 4: Proposed Information Extractor and Success Rate Predictor

5.2.2 Success Rate Predictor (SRP)

A legal case can be disputed and dragged for years to avoid this judge's success rate being computed. The case type information cannot be extracted by IE. Case type prediction is primarily based on the text in the judgment corpus. Keyword Extraction (KeyExt) is proposed to identify the keywords in the corpus based on a TextRank

approach. The identified keywords will be highly useful in finding the case type of the judgment corpus. It is a graph-based ranking algorithm that ranks text segment chunks in order of importance in the text document. The basic idea is to represent the relevant lexical unit as a graph and then ranking those chunks using the well-known PageRank algorithm. The words are scored in the judgment text and the keywords are found accordingly. The previously found case types and keywords will be stored in a database which will be used by the SRP to find the case type from the extracted keywords. The keywords will vary based on the case type. Initially manual extraction of keywords are performed based on the case types by legal experts. The SRP is trained with the labeled dataset. The process commences by training the model using the Bi-LSTM. The case type information is stored along with associated extracted features. This data is further used to compute the success rate of a judge with high accuracy.

5.2.3 Experimental Results

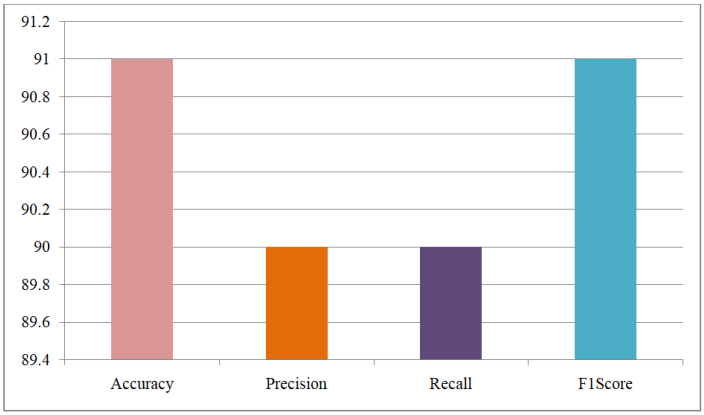

Figure 5: Evaluation results of element extraction in 2000 cases

During the evaluation phase, 2000 different case documents were randomly selected for testing. Figure 5 shows the results of the element extraction of 2000 cases. With an F1-score value of 0.90, precision of 0.90 and recall of 0.91, the prediction mechanism performed well while testing. The confusion matrix table for the testing and overall

datasets was calculated to assess the performance of our proposed method. Table 3 shows the accuracy of 94%.

Table 3: Cross-validation results for our records

Fold	Training set	Testing set	Accuracy (%)	Precision (%)	F1-score (%)	
1	8000	2000	94.25	94.09	94.5	94% overall accuracy
2	8000	2000	94.35	93.56	93.32	
3	8000	2000	94.32	91.86	92.38	
4	8000	2000	94.48	93.42	93.65	
5	8000	2000	94.03	93.61	90.5	

5.3 Legal Document summarization

A novel optimal Deep Learning based legal text summarization (ODL-ITS) technique has been designed to automatically generate the summary of the legal text documents. Figure 6 depicts the working of the proposed ODL-ITS. The technique contains a similarity measurement process, Bidirectional Gated Recurrent Neural Network (BiGRNN) based summary generation, and Glowworm Swarm Optimization (GSO) based hyperparameter tuning.

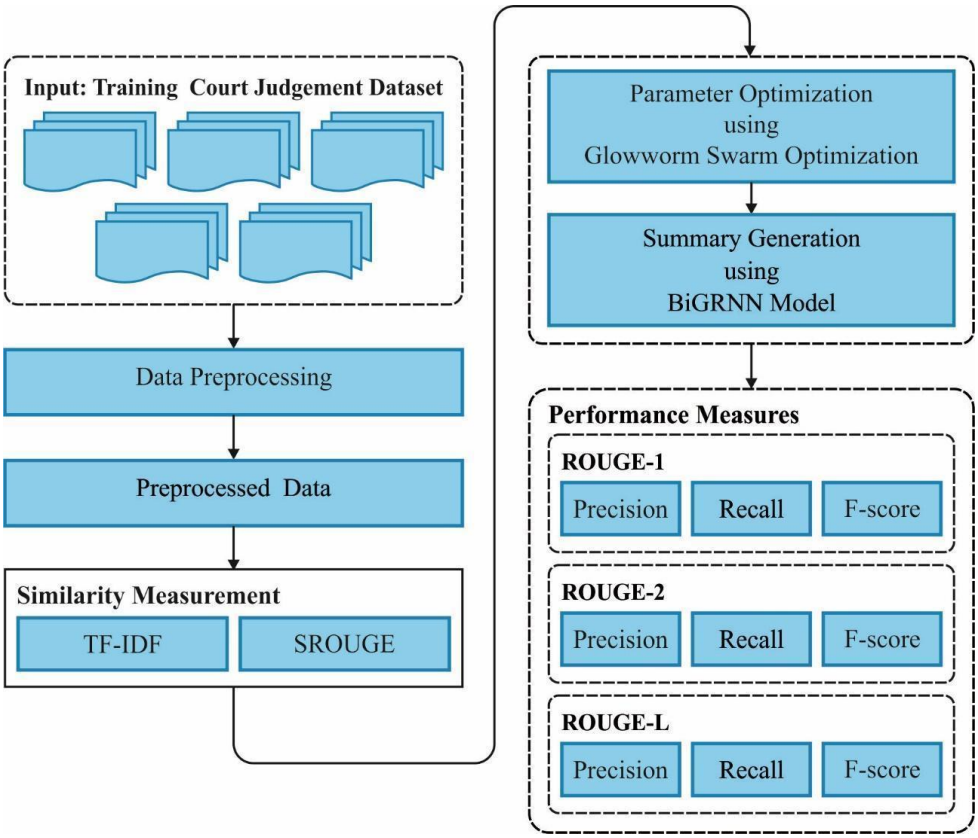

Figure 6: Overall process of ODL-ITS model

5.3.1. Similarity Measurement using TF-IDF

Term Frequency-Inverse Document Frequency (TF-IDF) explores important data in the corpus, and checks out related information to generate the summary in the document. The TF calculates the word in a document, thus the significant information for the summary is the word that appears prominently in the document. The IDF counts the number of times a word appears in a corpus of documents.

5.3.2. Similarity based on Rouge-L scores (SROUGE)

SROUGE is a set of metrics for assessing the supremacy of automated summaries, which is determined by comparing them against summaries provided by humans. The ROUGE-L metric belongs to the group of estimating methods that assess the superiority of a given

summary over a reference summary of the Longest Common subsequence (LCS). The ROUGE-L is defined as a fraction of text that overlaps the text in a produced summary where the text's granularity is considered to be a generic subsequence.

5.3.3. Summary Generation

The BiGRNN model is used in the summary generating process. The BiGRNN method treats summary generation as a classification problem, dividing a sentence into major and insignificant segments. Using a compositional BiGRNN, this model captures the context of the word in the sentences as well as the extended contexts between the sentences. It is made up of two layers of bidirectional GRU, the first of which converts sentences to feature vector representation and the second of which achieves document representation using this sentence vector. In this section, the two degrees of attention are employed to increase the weights of significant sentences and words.

The GSO technique is used to optimize the hyperparameters of the BiGRNN algorithm in order to increase its performance. It's a new type of metaheuristic and stochastic optimization method that was created by Indian researchers. GSO employs a swarm of glowworms as its agent, which is seen to be a potential solution to the challenges. The user-defined objective function is used to determine optimal fitness.

5.3.4 Experimental Analysis

The results showed that the ODL-LTS technique produced better summary generation results for various summaries sizes. Figure 7 shows the ODL-LTS technique's findings for rogue-1 and various summary sizes. In which ODL-LTS technique achieved the best summary generating results regardless of the summary size. For instance, with a summary size of 10%, the ODL-LTS technique has attained a precision of 38%, recall of 39.50%, and F-score of 38.50%. Besides, with a summary size of 50%, the ODL-LTS approach has gained a precision of 19.70%, recall of 67%, and F-score of 32%.

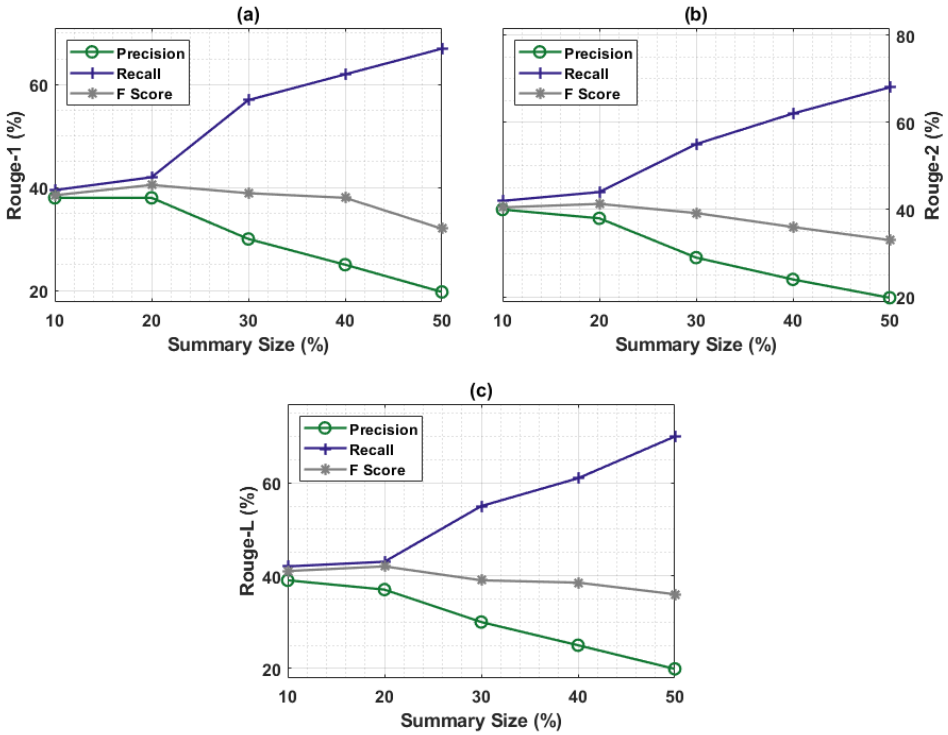

Figure 7: Result analysis of ODL-LTS model under different summary sizes

5.4 Deep Learning based Legal Data Classification

In the legal sphere, text classification is required to make reference easier for lawyers. The Judgment text is classified into eight different categories as shown in Figure 8. To accurately categorize the data, the words in the corpus file are transformed to vectors using word embedding techniques and the vectors are fed into the Long Short Term Memory for classification.

5.4.1 Word Embedding

The Continuous Bag of Word model (CBOW) is used for word embedding. CBOW predicts the target with the context word in the legal document. The context word and target word are used to transform the sentence into word pairs. The size of the context window is based on the document length. The words are vectored and projected to the

same positions using a projection layer, after which the vectors are averaged. Finally an element wise summation is performed on the vectors and output is obtained.

5.4.2 Convolution Neural Network and Long Short Term Memory

This phase uses the vectors obtained from CBOW and then features are extracted using Convolution Neural Network (CNN) further the features are optimized using Long Short Term Memory (LSTM). LSTM retains the long distance context information which is used for the precise classification of text. The LSTM has three gates. The first is the forget gate whose job is to decide which context information to be retained and which unnecessary data to be removed. Followed by the forget gate, there will be an input gate, which stores the new information. It has two layers, sigmoid layer and tanh layer. The sigmoid layer determines the information to be updated and the tanh layer creates new candidate information. In the output gate, the sigmoid layer finds the status of the output and tanh layer processes the output information.

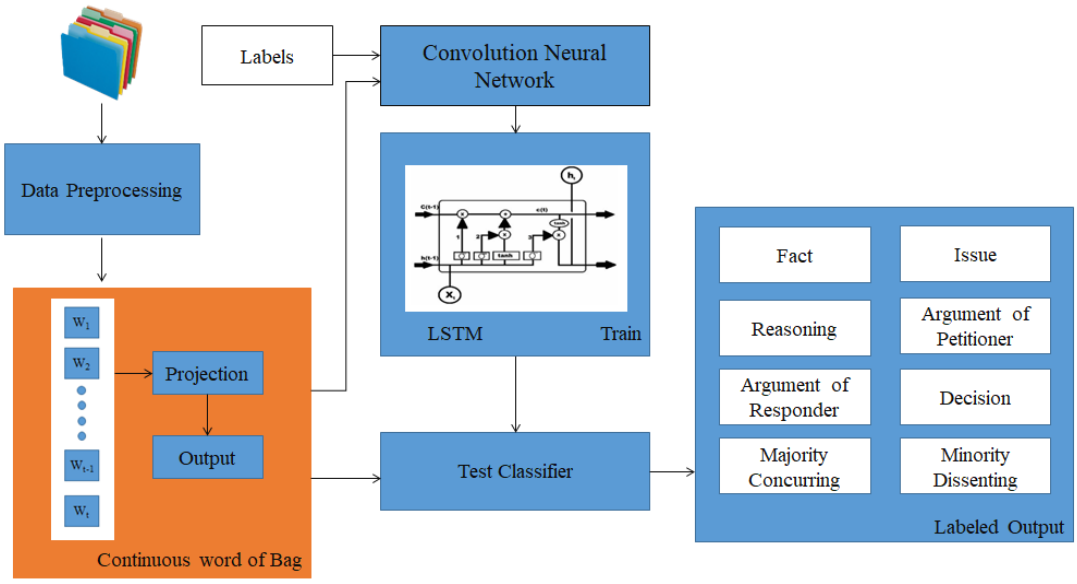

Figure 8: Legal document classification

5.4.3 Experimental analysis

The cleaned corpus file is split into the ratio of 90:10 for training and testing. When the model is tested, it classified the text with the accuracy of 0.95. For Facts, the average Issues is 0.93, and the medium Argument of Petitioner (AP) value is 0.96, the medium of Argument of the Responder (AR) is 0.94, Reasoning accuracy is 0.93 and the accuracy of decision is 0.98, the accuracy for Majority Concurring (MC) is 0.92 and Minority Dissenting (MD) accuracy is 0.93 as shown in Figure 9. The average accuracy of judgment text classification is higher than 90%. Initially it was classified to six labels Fact, Issue, Reasoning, Arguments of petitioner, arguments of responder, and Decision but there were misclassifications and the majority concurring and minority dissenting were wrongly placed into other categories. The Supreme Court lawyers suggested changing the six classes into eight class labels, which further improved the accuracy for the test data. Table 4 shows the comparison of proposed LDC with other models in which LDC outperforms the existing models.

Table 4: Comparison of LDC with Existing approaches

Model	Accuracy
Recurrent Neural Network	92.58
Hierarchical Attention Network	78.76
Deep CNN	79.28
Transformer	89.33
LDC (CNN+LSTM)	94.38

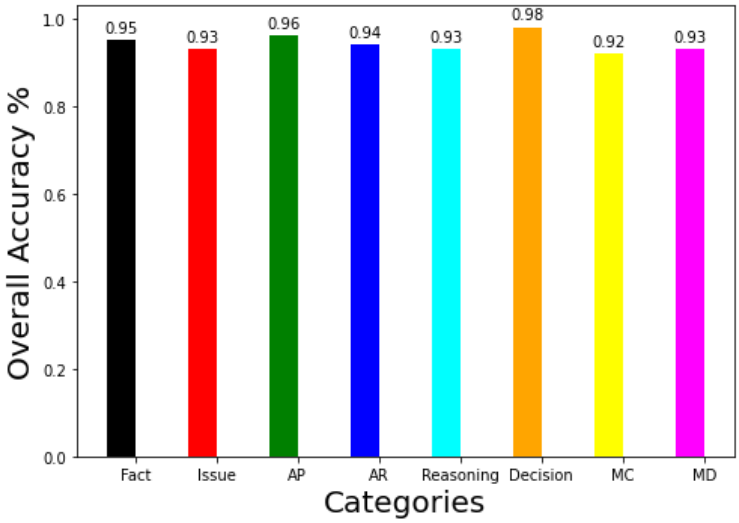

Figure 9: Accuracy of text classification based on categories.

6. ORGANIZATION OF THESIS

Chapter 1 - *"**Introduction**"* describes the basic terminologies in the legal domain. The current concepts adapted in the legal domain are discussed along with its advantages and disadvantages. The research motivation and objective to overcome the mentioned drawbacks along with the contribution and organization is elaborated.

Chapter 2 - *"**Literature Review**"* elaborates the current analytical research in legal domain and the associated existing techniques used for legal case outcome prediction, success rate prediction, legal document summarization and classification. Furthermore, the advantages and disadvantages associated with existing techniques are discussed.

Chapter 3 - *"**Design and Implementation of Deep learning based Opinion Mining Technique for Case Outcome Prediction**",* provides the Global Vector for Word Representation (GloVe) for converting legal data into vectors. Further CNN and hyper parameter tuning is explained to anticipate the outcome of the case file. Finally, comparison analysis of the proposed approach with the existing approaches is carried out.

Chapter 4 - *"**Design and Implementation of a Novel Keyword Extraction Technique and Success Rate Predictor for case type prediction**",* presents the detailed discussion of

novel key extraction. The Success Rate Predictor (SRP) is elaborated along with Bidirectional Long Short Term Memory (Bi-LSTM). Lastly, a comparison analysis of the proposed methods with the existing methods is performed.

Chapter 5 - *"**Design and Implementation of an Automated Optimal Deep Learning based Legal Text Summarization (ODL-LTS) Approach for legal text summarization**",* elucidates the similarity measure techniques like term frequency-inverse document frequency and ROUGE. Bidirectional Gated Recurrent Neural Network (BiGRNN) for summary generation is presented. Finally, in comparative analysis, the proposed methodology is compared to the existing methods.

Chapter 6 - *"**Design and Implementation of a Multiclass Text Classification Approach using CBOW and LSTM for Legal judgment text**"* describes CBOW and LSTM classification techniques. Lastly, the proposed approach is compared to the existing methods in a comparison analysis.

Chapter 7 - *"**Conclusion and Future work**"*, specifies the research contributions and future research perspectives.

REFERENCES

[1] Mehanna, Y. S., and Mahmuddin, M. B., A Semantic Conceptualization Using Tagged Bag-of-Concepts for Sentiment Analysis, *in IEEE Access*, vol. 9, 2021, 118736-118756,.

[2] Deng, D., Jing, L., Yu, J., and Sun, S., Sparse Self-Attention LSTM for Sentiment Lexicon Construction, *in IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 11, Nov. 2019, 1777-1790.

[3] Habibi, M., and Popescu-Belis, A., Keyword Extraction and Clustering for Document Recommendation in Conversations, *in IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 4, April 2015, pp. 746-759.

[4] Akhter, M. P., Jiangbin, Z., Naqvi, I. R., Abdelmajeed, M., Mehmood, A., and Sadiq, M. T., Document-Level Text Classification Using Single-Layer Multisize Filters Convolutional Neural Network, *in IEEE Access*, vol. 8, 2020, pp. 42689-42707.

[5] Li, G., Wang, Z., and Ma, Y., Combining Domain Knowledge Extraction With Graph Long Short-Term Memory for Learning Classification of Chinese Legal Documents, *in IEEE Access*, vol. 7, 2019, 139616-139627.

[6] Ghodratnama, S., Beheshti, A., Zakershahrak, M., and Sobhanmanesh, F., Extractive Document Summarization Based on Dynamic Feature Space Mapping, *in IEEE Access*, vol. 8, 2020, pp. 139084-139095.

[7] Gidiotis, A., and Tsoumakas, G., A Divide-and-Conquer Approach to the Summarization of Long Documents, *in IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, 2020, pp. 3029-3040.

[8] Jianqiang, Z., Xiaolin, G., and Xuejun, Z., Deep convolution neural networks for Twitter sentiment analysis, *IEEE Access*, vol. 6, (2018), pp. 23253–23260.

[9] Hyun, D., Park, C., Yang, M.-C., Song, I., Lee, J.-T., and Yu, H., Target-aware convolutional neural network for target-level sentiment analysis, *Inf. Sci.*, vol. 491, Jul.2019, pp. 166–178.

[10] Ma, Y., Peng, H., Khan, T., Cambria, E., and Hussain, A., Sentic LSTM: A hybrid network for targeted aspect- based sentiment analysis, *Cogn. Comput.*, vol. 10, no. 4, Aug.2018, pp. 639–650.

[11] Hu, F., Li, L., Zhang, Z.-L., Wang, J.-Y., and Xu, X.-F., Emphasizing essential words for sentiment classification based on recurrent neural networks, *J. Comput. Sci. Technol.*, vol. 32, no. 4, Jul.2017, pp. 785–795.

[12] Karim, R., Islam, M.A., Simanto, S.R., Chowdhury, S.A., Roy, K., Al Neon, A., Hasan, M., Firoze, A. and Rahman, R.M., A step towards information extraction: Named entity recognition in Bangla using deep learning, *Journal of Intelligent & Fuzzy Systems*, 37(6), 2019, pp.7401-7413.

[13] Li, J., Zhang, G., Yu, L. and Meng, T., Research and design on cognitive computing framework for predicting judicial decisions, *Journal of Signal Processing Systems*, 91(10), 2019, pp.1159-1167.

[14] Liu, Y.H., and Chen, Y.L., A two-phase sentiment analysis approach for judgment prediction, *Journal of Information Science*, 44(5), 2018, pp.594-607.

[15] FirdausSolihin, Indra Budi, Rizal FathoniAji&EdmonMakarim, Advancement of information extraction use in legal documents, *International Review of Law, Computers & Technology*, 35:3, , 2021, 322-351.

[16] D. Anand., and R. Wagh., Effective deep learning approaches for summarization of legal texts, *Journal of King Saud University-Computer and Information Sciences*, 2019.

[17] Tiwari, A., and Dembla, D., A Novel Algorithm for Automatic Text Summarization System Using Lexical Chain, *In Ambient Communications and Computer Systems*, 2019, pp. 103-112.

[18] Zhang, Y., Li, D., Wang, Y., Fang, Y. and Xiao, W., Abstract text summarization with a convolutional Seq2seq model, *Applied Sciences*, 9(8), 2019, p.1665.

[19] Rahman, M., and Siddiqui, F.H., An optimized abstractive text summarization model using peephole convolutional LSTM, *Symmetry*, 11(10), 2019, p.1290.

[20] Wan, L., Papageorgiou, G., Seddon, M., &Bernardoni, M. Long-length Legal Document Classification, *arXiv preprint arXiv:*1912.06905, 2019.

[21] Iftikhar, A., Jaffry, S. W. U. Q., & Malik, M. K., Information mining from criminal judgments of Lahore high court, *IEEE Access : Practical Innovations, Open Solutions*, 7, 2019, 59539–59547.

[22] Yan, G., Li, Y., Shen, S., Zhang, S., & Liu, J. (2019, July). Law Article Prediction Based on Deep Learning, *In 2019 IEEE 19th International Conference on Software Quality, Reliability, and Security Companion* (QRS-C), July 2019, pp. 281-284.

LIST OF PUBLICATIONS

Journal Publications

- Vaissnave, V., and Deepalakshmi, P., Data Transcription for India's Supreme Court Documents Using Deep Learning Algorithms', International Journal of Electronic Government Research (IJEGR), 16(4), (2020), 21-41.
- Vaissnave, V., and Deepalakshmi, P., Analysis on Hybrid Deep Neural Networks for Legal Domain Multitasks: Categorization, Extraction, and Prediction, International Journal of e-Collaboration (IJeC), 18(1), (2022), 1-22.
- Vaissnave, V., and Deepalakshmi, P., Modeling of Automated Glowworm Swarm Optimization Based Deep Learning Model for legal text summarization, International Journal of Multimedia Tools and Applications. (Accepted) IF: 2.5
- Vaissnave, V., and Deepalakshmi, P., A Threshold-based POS Tagging and Gradual Magnitude Pruning Compression Model using Keyword Extraction and Case Type Prediction on the Supreme Court of India Data, International Journal of Intelligent and Fuzzy Systems.(IF:1.7). (Received Revisions).

Conference Publications

- Vaissnave, V., and Deepalakshmi, P., An Artificial Intelligence based Analysis in Legal domain International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-2S2, December 2019.
- Vaissnave, V., and Deepalakshmi, P., A Keyword-Based Multi-label Text Categorization in the Indian Legal Domain Using Bi-LSTM. In Soft Computing: Theories and Applications, 2021, (pp. 213-227). Springer, Singapore.
- Vaissnave, V., and Deepalakshmi, P., Sentiment Analysis for Legal Judgment Text in India’s Supreme Court Based on Glove Pre-trained Word Embedding and Deep Learning Models, Soft Computing: Theories and Applications, 2022, (pp. 213-227). Springer, Singapore. (SCOPUS).